DOCUMENT RESUME

ED 214 259 EA 014 485

AUTHOR Seidman, William
TITLE Goal Ambiguity and Organizational Decoupling: The Failure of Rational Program Implementation.
PUB DATE Mar 82
NOTE 49p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, March 19-23, 1982).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Compliance (Legal); *Educational Change; Educational Legislation; Educational Objectives; Elementary Secondary Education; Evaluation Criteria; *Federal Programs; *Federal Regulation; Government School Relationship; National Programs; Outcomes of Education; *Program Evaluation; *Program Implementation; Social Problems; State Programs
IDENTIFIERS Elementary Secondary Education Act Title I

ABSTRACT

Noting a general dissatisfaction with the implementation and evaluation of large-scale educational programs, this paper argues that the rational model that is the foundation for large-scale educational programs is faulty. According to the rational model, government mandates call for educational programs to remedy social problems; then rewards and sanctions (applied according to the results of program evaluations) force schools to implement these programs; and finally the resulting changes alleviate the original social problems. One difficulty with the model is that there is often a tenuous link between actual program intent and evaluative criteria. Evaluative criteria chosen are usually those easily measurable, like test performance, rather than those that might actually indicate that the rather ambiguous program goals (such as reduction of poverty) have been realized. Thus, program intent may at times be decoupled from the evaluative criteria used to force program implementation. A second problem concerns the many possible focuses of evaluation: program compliance, program implementation, educational outcomes, or social outcomes. It has not been demonstrated that evaluations of any of these factors are effective in achieving program intent. For this reason, schools often decouple such dubious outcome data from program decisions. (Author/JM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED214259

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

William Seidman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

GOAL AMBIGUITY AND ORGANIZATIONAL DECOUPLING: THE FAILURE OF RATIONAL PROGRAM IMPLEMENTATION

William Seidman

March, 1982

EA 014 485

Presented at the American Educational Research Association Annual Meeting, New York, NY.

There has been, in recent years, a growing dissatisfaction with large scale education programs and with the evaluation of these programs. Decision makers feel that programs have not been implemented as intended and that program evaluations have provided little useful information. Evaluators feel that both these decision makers and the public ignore or misinterpret their findings. Frustration with large-scale education programs and program evaluation abounds.

This paper examines the implementation and evaluation of large-scale educational programs. Specifically, this paper argues that the rational model of government control of schools that is the foundation for educational programs and program evaluations is faulty. Consequently, expectations based on this model are not met and dissatisfaction grows.

Background

Since the early 1960's large-scale federal and state programs -- designed to alleviate social problems (Wildavsky, 1979) -- have been a prominent feature of American education. Such programs as Title I of the Elementary and Secondary Education Act and California's School Improvement Program were (and still are) expected to improve the education of disadvantaged students or increase parent participation in school decision making. A plethora of government-run social programs is expected to reduce social problems by improving the schools.

The assumption that governmental action can improve schools is based on a rational model of program implementation and evaluation (Figure 1) (House, 1979; Kagan, 1981; Sproul, forthcoming; Wise, 1979).

The generic form of this model has five components:

1. Legislation is passed or judicial decisions are handed down that call for programs to remedy social problems;
2. Federal or state offices or departments of education:
   a. convert the legislator's intent into specific program guidelines;
   b. develop evaluative criteria for examining the program's effectiveness; and
   c. administer rewards and sanctions to schools and school districts according to local success in the evaluation.
3. Driven by the rewards and sanctions to improve performance on the evaluative criteria, schools change instructional and other practices.
4. These new practices produce improvements in particular educational outcomes;
5. Improved educational outcomes lead to social outcomes that remedy the original social problem.

Rewards and sanctions -- applied by educational bureaucrats according to the results of program evaluations -- are assumed to pressure schools into changing their practices (Berman, 1978, 1981; Elmore, 1980; Kagan, 1981; Sabatier and Mazmanian, n.d.; Sproul, 1981a; Wildavsky, 1979). These changes in local practice are expected to produce desired educational outcomes which are, in turn, expected to alleviate the original social problem (Berman, 1978; Wildavsky, 1979).

A hypothetical vocational education law illustrates this model. Concerned about unemployment among certain groups of minorities, Congress

passes a law calling for and funding innovative, targeted, vocational educational programs. The Department of Education defines "innovative," develops formulas for insuring that the funding is targeted to minorities, evaluates the program according to the definition of innovative and the correctness of the funding formula, and on the basis of this evaluation, continues or discontinues the funding. Schools, driven by the evaluation and the threat of decreased funding, change their programs to meet the evaluative criteria. In doing so, these schools teach minority students new vocational skills that make them more attractive employees, thereby increasing both their employment opportunities and, presumably, their actual employment. Programmatic change is caused by evaluative pressure and leads to the remediation of a social problem. The rational and bureaucratic model of changes is assumed to quickly and efficiently produce desired social outcomes (McLaughlin, 1975).

Although the model presented above may appear simplistic, several authors (National Institute of Education, 1977; Sproul, 1981a) imply that just such a simplified model underlies most program implementation efforts. Further, the progression from legislative or judicial initiative through bureaucratic and local responses to educational and social outcomes is described in essentially this simplistic form by many authors (see, for example, Benson, 1980; Elmore, 1980; House, 1979; Kagan, 1981; Kirst and Jung, 1981; National Institute of Education, 1977; Sabatier and Mazmanian, n.d.; Sproul, 1981a, forthcoming; Wilkes, Eaton, Wood and Dornbusch, 1979; Wise, 1979). The assumed progression of reform initiatives from legislative intent to social outcomes is well supported. However, the notion that

evaluation is a control mechanism rather than a means of informing policy makers (Elmore, 1980) is just now receiving attention and, therefore, requires some additional examination.

In their research on evaluation, Dornbusch and Scott (1975) found that evaluation, though influenced by many variables, tends to control people's behavior. When people are evaluated according to identified criteria, and when rewards and sanctions are dispensed according to the results of these evaluations, evaluatees attempt to improve their performance on the evaluated criteria. For example, if a production worker is evaluated and paid according to the number of units produced, this worker is likely to try to produce more units. In order to gain rewards or avoid sanctions, people work to improve performance on the evaluative criteria.

Policy makers and evaluators have made a similar assumption about the effects of program evaluation on school behavior since 1965. At that time, both Senator Robert Kennedy and Commissioner of Education Francis Keppel indicated that evaluation of Title I was needed to ensure the proper use of Title I funds (Elmore and McLaughlin, 1981; McLaughlin, 1975; Wildavsky, 1979). These Title I evaluations were to focus parental attention on the conditions in their children's schools, thereby making schools both more accountable to parents and more responsive to change initiatives. Essentially, program evaluations were seen as a necessary means of forcing program implementation (Elmore and McLaughlin, 1981; McLaughlin, 1975). Without such force, it was assumed, programs would not be implemented.

More recently, many authors (Elmore, 1980; Elmore and McLaughlin, 1981; House, 1979; Kirst and Jung, 1981; McLaughlin, 1975; Rutherford and

Hoffman, 1981; Sabatier and Mazmanian, n.d.; Wilkes, Eaton, Wood and Dornbusch, 1979; Wildavsky, 1979) have advanced a similar argument. These authors contend that program evaluations that have clear evaluative criteria and firmly applied sanctions can (and have) forced program implementation. For example, Kirst and Jung (1981) reported that increased targeting of Title I funds to appropriate student groups is a result of tighter evaluations and firmer enforcement. Program evaluation is still seen as a means of forcing implementation of educational reforms.

Furthermore, school administrators and teachers believe that program evaluations are a means of controlling their behavior (David, 1981; Rutherford and Hoffman, 1981; Sproul, 1981a). Program evaluations are perceived as threatening (David, 1981; Sproul, 1981a) in general and as a threat to resources (Rutherford and Hoffman, 1981) in particular. Since poor evaluations are seen as harmful to schools, schools can be expected to attempt to improve their evaluations by conforming to the evaluative criteria. Schools view program evaluations more as an "instrument of control" (Elmore and McLaughlin, 1981, page 2) than as a means of informing policy.

In summary, while few educators, administrators or policy makers would argue that the rational model of educational change captures the complexity of the educational change process, this model seems to capture many of the assumptions underlying government efforts to change the schools. Both the legislative intent to social solutions progression and the use of program evaluation as a control mechanism are well supported by the program implementation and evaluation literature.

Yet, programs implemented and evaluated using this model have not produced the anticipated results (Berman, 1978; Cronbach, 1980; Majone, 1980; Wholey, 1981; Wildavsky, 1979; Wise, 1979). People complain that schools do not implement programs (Berman, 1978; Majone, 1980; Wholey, 1981). Decision makers complain that program evaluations have neither forced program implementation nor produced useful data for program revisions (Berman, 1978; Majone, 1980; Wildavsky, 1979). Evaluators complain that neither the schools nor the politicians use their findings (Alkin, Daillak and White, 1979; Cronbach, 1980; Wholey, 1981). The rational model of program implementation and evaluation seems to have failed.

This paper argues that the failure of large scale, government run education programs is due, at least in part, to weaknesses in the rational model. Specifically, this paper argues (1) that the evaluative criteria found in most program evaluations is only loosely-coupled to the program's original intent, and (2) that evaluative pressure does not produce the anticipated site-level responses. This paper does not directly examine the effects of implementation on educational outcomes or of educational outcomes on social outcomes because these topics are too broad to be thoroughly examined in the time and space available, and because low levels of program implementation decrease the necessity of examining these linkages.

Definitions

Before proceeding, it is necessary to define both the scope of this analysis and some of the terminology used in this paper. This paper examines only federal and state efforts to change schools. Thus, only large scale programs and their evaluations are analyzed while local initiatives and

evaluations are ignored. This limitation focuses attention more directly on the rational model.

The terms evaluative criteria and evaluative pressure appear frequently. Evaluative criteria are the specific attributes of a program that are measured for the purposes of determining the distribution of rewards and sanctions (Dornbusch and Scott, 1975). For example, conformity to the Individualized Educational Program (IEP) requirements in PL94-142 is one of the evaluative criteria for PL94-142 since funding is, in theory, conditional on meeting the IEP regulations. Evaluative criteria are the measured attributes of a program. Other common evaluative criteria include test scores and earnings. Evaluative pressure is the application of rewards and sanctions based on success or failure on the evaluative criteria. In PL94-142 evaluative pressure is exerted on compliance with the IEP requirement since funding is conditional on conformity to this requirement. Evaluative pressure directs attention to the evaluative criteria.

Compliance, implementation and educational and social outcomes also require definition. For the purposes of this paper, compliance is defined as conformity to pre-specified procedural guidelines (i.e. rules and regulations). Developing a formula for distributing vocational education monies is a compliant behavior for the Vocation Education Act of 1976, since the rules and regulations for that act require a state to have such a formula (Hoachlander, 1981). Compliance means conformity to procedural evaluative criteria. Implementation is the "translation of project plans and proposals into [classroom] practice" (Berman and McLaughlin, 1978, p. 15). The changes in program practice in Alum Rock (Weiler, 1976) constitute program

implementations since these changes generally followed the project's intent. Implementation, then, is actual program change that is in accord with program intent. As may be apparent, there may not always be a distinction between compliance and implementation since a compliant behavior may affect classroom behaviors. Educational outcomes are the "immediate" effects of schooling (Levin, 1974, page 86). Achievement test scores or particular vocational skills are educational outcomes since they exist at the immediate conclusion of a student's educational experience. In contrast, social outcomes are changes in the society. Less unemployment and reduced poverty are examples of social outcomes since they are changes in the society at large. Each of these components of the rational model has been used as evaluative criteria to force program reform.

Program Intent and Evaluative Criteria

The rational model of educational change assumes that the evaluative criteria used to force a program's implementation are highly congruent with a program's original intent (Berman, 1981; Majone, 1980; National Institute of Education, 1977; Sproul, 1981a; Wildavsky, 1979). However, this assumption is suspect (Berman and McLaughlin, 1978; House, 1979; Wildavsky, 1979). Because program intent is often ambiguous and social outcomes are frequently difficult to measure, evaluative criteria have typically been what is measurable, and what is measurable often has little relation to a program's intent (Cronbach, 1980; House, 1979; Kagan, 1981; Wildavsky, 1979; Wise, 1979). Consequently, program intent may, at times, be decoupled from the evaluative criteria used to force program implementation.

According to Schultze (in Cronbach, 1980, page 130) "The first rule of the successful political process is 'Don't force a specification of goals and ends." This rule serves the politicians' purposes by limiting political accountability and by allowing the appearance of noble ambitions, acceptable to broad constituencies. But it results in legislative language that is, at times, vague, ambiguous, or conflicting (Benveniste, 1981; Berman, 1978; McLaughlin, 1975; Sabatier and Mazmanian, n.d.; Wildavsky, 1979; Wise, 1979). For example, Hoachlander (1981) reports that the Vocational Education Act of 1976 called for programs aimed at "economically depressed areas" without providing any indication of how this should be defined. Similarly, PL94-142 calls for a "free and appropriate education" (FAPE) for all handicapped children, that is to be given in the "least restrictive environment" (Benveniste, 1981; Kagan, 1981; Stearns, et al., 1980). And Title I was purposely left vague in order to generate broader political support for its passage (Berman, 1978; Kirst and Jung, 1981; McLaughlin, 1975). Many educational initiatives are written in broad, idealized language in order to attract support (Wise, 1979).

Yet, "legislators who are quite sophisticated about their own reasons for keeping 'goals' lofty and nebulous unblushingly ask program administrators to make goals explicit" (Cronbach, 1980; pages 130-131). The California competency testing law, for instance, contains powerful statements about basic skills and requirements for certification of minimum competency but leaves the definition of minimum competency and of certification up to the State Department of Education and the local school districts. In much of the same manner, the Bilingual Education Act (ESEA Title VII) and the Lau

v. Nichols decision in bilingual-bicultural education called on the Office of Civil Rights in USOE to establish guidelines for implementing bilingual-bicultural programs (Berman, 1981). Program administrators and evaluators are forced to interpret indistinct, hazy social ideals and to convert these ideals into specific program regulations and well-defined evaluative criteria (Wise, 1979).

However, in the interpretation and conversion process differing interpretations of the meaning of legislation proliferate. PL94-142, for instance, can be interpreted either as an education law or as a civil rights law (Stearns, et al., 1980). The Bilingual Education Act was initially interpreted as calling for just bilingual programs but was reinterpreted in Lau v. Nichols as requiring bilingual-bicultural programs (Berman, 1981). Title I has been interpreted as requiring both strict and flexible enforcement of the funding regulations (National Institute of Education, 1977). Reasonable, but different, interpretations of legislative and judicial intent abound. As a result, it is difficult to determine if evaluative criteria are congruent with a program's intent, since there is often substantial disagreement about what a program's intent actually is.

In spite of the difficulties of determining a program's "true" intent, some evidence suggests that evaluative criteria are reduced to what is measurable and can be implemented, regardless of program intent (Cronbach, 1980; Elmore, 1980; House, 1979; Kagan, 1981; Wildavsky, 1979; Wise, 1979). For example, according to House (1979):

> Follow Through began with a broad set of social goals that it sought to accomplish . . . but the evaluators and the Office of Education ultimately succeeded in reducing the outcomes to a few common measures

on which scaled comparisons could be made (p. 35).

The broad social goals of Follow Through -- and of Title I, too (Cronbach, 1980) -- were reduced to "trivialized" (Wise, 1979: page 5) evaluative criteria, standardized tests (House, 1979). Similarly, although the goal of the Alum Rock voucher experiment was improved instruction, Weiler's (1976) evaluation actually measured diversity of instruction. Measurable criteria were substituted for an undefinable goal. Other programs that exhibit this displacement between program intent and evaluative criteria include: Chapter 766, the Massachusetts "mainstreaming" law (Weatherley and Lipsky, 1976), PL94-142 (Stearns, et al., 1970) and the Vocational Education Act of 1976 (Hoachlander, 1981). In each of these programs broad social goals are replaced by narrower, measurable criteria that may be only weakly linked to the original intent. Evaluative criteria are, at times, only loosely coupled to program intent.

Unfortunately, while the processes of goal displacement have been examined in the organizational behavior literature -- see, for example, Dornbusch and Scott (1975), March and Olsen (1976), Pfeffer and Salancik (1978), Scott (1981), and Weick (1979) -- there have been few, if any, studies of the relationship between program intent and evaluative criteria in large-scale educational programs. Therefore, although the evidence for loose coupling between intent and evaluative criteria is strong, it is not conclusive.

Nevertheless, the evidence suggests that the rational model's assumption that evaluative criteria will be congruent with program intent may

not always be correct. Evaluative criteria may, at times, have little relationship to intent. Consequently, even if implementation were to occur and evaluation was effective, the program would be seen as a poor response to that problem and evaluation of the program would be seen as producing inappropriate data. The decoupling of this section of the rational model is likely to cause dissatisfaction among decision makers and evaluators alike.

Evaluation as a Control Mechanism

In addition to the assumed linkage between program intent and evaluative criteria, the rational model assumes that evaluative pressure can force schools and school districts to behave in desired ways (Berman, 1978, 1981; Elmore, 1979, 1980; Kagan, 1981; Sabatier and Mazmanian, n.d.; Sproul, 1981a). However, a theoretical analysis of this assumption and a review of the program implementation and evaluation literature suggest that there are actually four variations of this assumption (Figure 2). These variations -- are distinguished by the focus of evaluative pressure (on compliance and implementation processes or on educational and social outcomes), by the amount of evaluative pressure applied to schools, by the degree of autonomy allowed schools in responding to the pressure and by the different site-level reactions the evaluative pressure causes. None of the site-level responses caused by the evaluative pressure conform to the rational model.

Evaluating Compliance

Evaluations of compliance (Figure 2a) -- the most common form of evaluative pressure exerted on schools (Elmore and McLaughlin, 1981; Sproul, 1981a) -- are expected to force schools to closely follow, or comply with, program rules and regulations. Compliance, it is assumed, is either

synonymous with or will force implementation. Implementation is expected to lead first to desirable educational outcomes then to appropriate social outcomes. In this regulatory (Elmore, 1979) or compliance variation, program evaluation of compliance forces educational change (Berman, 1981; Elmore, 1980; Elmore and McLaughlin, 1981; Kirst and Jung, 1981; Sabatier and Mazmanian, n.d.; Sproul, forthcoming; Wilkes, Eaton, Wood and Dornbusch, 1979; Wildavsky, 1979).

Although PL94-142 does not explicitly state its social objectives, several authors (Beneviste, 1981; Elmore and McLaughlin, 1981; Sproul, 1981a; Stearns, et al., 1981) consider it a particularly good illustration of compliance approach to program implementation. In PL94-142 evaluative pressure is focused, to a great degree, on the IEP requirement. This pressure is expected to force school districts to write IEP's for every handicapped child in the district. The process of writing an IEP, it is assumed, will automatically insure that students receive a "free and appropriate education," the goal of the law (Elmore and McLaughlin, 1981; Stearns, et al., 1980). Evaluative pressure exerted on compliance with rules and regulations is expected to force compliance, and compliance is expected to produce, eventually, desired social outcomes.

It is important to emphasize that the purpose of this approach is to force local obedience to government policy by restricting local behavior (Elmore, 1980; Elmore and McLaughlin, 1981; Gideonse, 1980; Kirst, 1981; National Institute of Education, 1977; Shapiro, 1980; Wildavsky, 1979). Such a restrictive approach is based on the assumption that government knows more about education than local educators (Shapiro, 1980), and that local

educators can not be trusted to implement a program (Elmore, 1980; Gideonse, 1980; Kirst, 1981; National Institute of Education, 1977). Proscribing local discretion and initiative is seen as a necessary pre-requisite to forcing program implementation.

Yet, evaluation of compliance does not always produce the anticipated results (Berman, 1981; Wildavsky, 1979). While it is fairly well established that evaluation of compliance typically produces compliance (Berman, 1981; Elmore, 1980; Elmore and McLaughlin, 1981; Kagan, 1981; Kirst and Jung, 1981; National Institute of Education, 1977; Weatherley and Lipsky, 1976; Wilkes, Eaton, Wood and Dornbusch, 1979), compliance does not necessarily lead to implementation. Rather, evaluation of compliance has at least three unexpected results, each of which may be, at times, counter-productive (Berman, 1978, 1981; Elmore, 1979, 1981; Kagan, 1981; Wildavsky, 1979; Wise, 1979).

The evidence supporting the notion that evaluation of compliance produces compliance is fairly strong. Berman (1981), for example, reports that the Los Angeles school district developed a massive in-service training program in foreign languages and cultures in order to have enough bilingual teachers to meet the requirements of the Lau v. Nichols bilingual-bicultural decision. Similarly, studies of two "mainstreaming" laws, Massachusett's Chapter 766 (Weatherley and Lipsky, 1976) and PL94-142 (Hassel, in Berman, 1981; Stearns, et al., 1980) indicate that "the threat of . . . monitoring helped ensure the compliance of local special education administrators" (Weatherley and Lipsky, 1976, page 179) to the laws' rules and regulations. And Kirst and Jung (1980) report that increased accuracy of

the targeting of Title I funds to intended student populations was a direct result of frequent evaluations of compliance and of firm administration of sanctions based on these evaluations. In sum, these and other studies (e.g., Felix, 1979; Hoachlander, 1981; Sproul, 1981a, forthcoming) clearly support the rational assumption that evaluation of compliance produces compliance.

But compliance may not always produce implementation and may actually retard implementation (Berman, 1978, 1981; Elmore, 1979, 1981; Kagan, 1981; Wildavsky, 1979; Wise, 1979). Specifically, a concern for compliance "leads to a considerable amount of nonproductive effort [mostly paperwork and meetings], diverting those local educators who are in fact dedicated to the statutory goals from important tasks to low priority 'compliance' activities" (Kagan, 1981, page 52). For instance, in their study of the California Early Childhood Education (ECE) Program (which had highly specific evaluative criteria that were designed to force improvements in basic skills education) Cohen and Miller (1980) found that ECE schools developed more coordination (number of policies and number of committees) than non-ECE schools but that the ECE principals evaluated their teachers less than non-ECE principals. Formalized structures that complied with the evaluative criteria increased, but control of implementation of the program diminished.

Virtually identical findings are reported by Weatherley and Lipsky (1976) for Chapter 766 and by Hassel (in Berman, 1981) and Stearns, et al. (1980) for PL94-142. Schools in these programs were often more concerned with "completing forms properly and speedily" (Weatherley and Lipsky, 1976, page 185) than with implementing the program. Furthermore, time needed to coordinate IEP meetings (termed "cores" in Massachusetts) and write IEP's was

"legendary" (Kagan, 1981, page 38), and usurped time that was better spent with children. Compliance activities are often a "diversion of energy from more important matters," (Sproul, 1981a: page 7) attention to children and program implementation. Forced compliance may, at times, decrease implementation (Alkin and Daillak, 1979; Berman, 1978, 1981; David, 1981; Elmore, 1980; Elmore and McLaughlin, 1981; Hoachlander, 1981; Kagan, 1981; National Institute of Education, 1977; Sproul, 1981a; Wildavsky, 1977). The rational assumption that compliance leads to implementation is sometimes incorrect.

However, the counterproductive effects of focusing evaluative pressure on compliance are not limited to a decrease in attention. In some circumstances, concern for compliance leads to intentional and unintentional behaviors that sabotage a program's intent. For example, Hoachlander (1981) suggests that many of the funding formulas developed to comply with the regulations of the Vocational Education Act of 1976 were consciously designed to funnel money to wealthy districts rather than to the poor districts that were the intended targets of the law. Similarly, Monti (1979) reports that a Missouri school district's compliance with a desegregation order led both to the exclusion of many formerly active black parents from district decision making and to a decrease in the percentage of black faculty in the district. Also, Beneviste (1981) and Stearns, et al. (1980) indicate that relatively wealthy parents of handicapped children have used the PL94-142 appeals procedure to gain special placements for their children at public expense, thereby clogging the appeals channels and causing districts to recommend only inexpensive IEP's. Concern for compliance can lead to or allow school

responses that are counter to a program's intent (Benson, 1980; Berman, 1981; Elmore and McLaughlin, 1981; Gideonse, 1980; House, 1979; Kagan, 1981). Adherence to the rational model be counterproductive.

Evaluation of compliance may be counterproductive in yet another way. Schools that attempt to implement a program's intent, without attending to compliance, sometimes receive poorer evaluations than schools that ignore implementation in favor of compliance (Berman, 1981; Hoachlander, 1981; Kagan, 1981). For example, Kagan (1981) reports that a New York City special education program that a federal judge thought was an "especially dedicated, expensive and substantial one" (page 40) was cited by HEW as being out of compliance, while Berman (1981) indicates that both Title I programs and bilingual programs that implemented the intent of these programs were often cited for non-compliance. Schools that try to create effective programs may be penalized. In turn, these unjust evaluations may discourage schools from attempting to implement a program's intent. Here too, evaluating compliance is likely to decrease implementation. The rational assumption that compliance causes increases in implementation is problematic.

There are many possible explanations for the failure of this variation of the rational model. Several theorists (Berman, 1981; Elmore and McLaughlin, 1981; Kagan, 1981; Kirst and Jung, 1981; Sabatier and Mazmanian, n.d.; Sproul, 1981a) argue that the decoupling of implementation from compliance is a result of weak enforcement of unclear guidelines. Clearer program objectives, better evaluations and firmer applications of sanctions, they suggest, will increase compliance and increased compliance will lead to program implementation. However, this perspective is still based on the

rational model, which is clearly suspect. Other authors (Stearns, et al., 1980; Weatherley and Lipsky, 1976) suggest that this decoupling is really a resource problem; that is, given limited resources, districts can either comply or implement, but not both. Since evaluative pressure is exerted on compliance, districts comply, to the detriment of implementation. Still other theorists suggest that the unknown technology of education (Berman, 1981; Elmore, 1981), complex and confusing rules and regulations (Sproul, 1981a), administrative communications patterns (Sproul, 1981b; Talbert, 1980) or a failure to operationalize program intent (Hoachlander, 1981) allows "standard operating procedures" (Cyert and March, 1963) or their equivalent to dominate school level responses to program initiatives (Berman, 1978; Sproul, 1981a, 1981b, forthcoming); that is, schools respond to reform efforts by integrating new guidelines into their normal functional procedures with as little disruption of routines as possible. Finally, schools may be "loosely-coupled" (Weick 1976) or "de-coupled" (Meyer and Rowan, 1977, 1978) organizations, responding to external pressures with administrative behaviors that are "buffered" (Meyer and Rowan, 1977, 1978) from substantive program change. While each of these last three arguments seems plausible, it is not yet possible (or appropriate) to identify one explanation as the "correct" one.

In summary, although it is not now possible to specify why compliance does not cause implementation, it is clear that the rational assumption that evaluation of compliance leads to program implementation is not always correct. Exerting evaluative pressure on compliance seems to produce compliance, but compliant behaviors may, at times, be counterproductive to

program implementation (Berman, 1981; David, 1981; Elmore, 1980; Elmore and McLaughlin, 1981; Kagan, 1981; Wildavsky, 1979). Thus, the validity of the rational model is suspect.

Evaluating Implementation

Evaluations of implementation (Figure 2b) present a somewhat different set of challenges to the rational model. Some evidence suggests that evaluations of implementation -- which are distinguished by an absence of evaluative pressure -- allow mutual trust to develop between evaluators and implementer, thereby promoting program implementation. But a closer analysis of this evidence indicates that informal evaluative pressure may still be exerted on implementation, accounting for some of the implementation, and that the program changes reported were in relatively unimportant projects. Nevertheless, implementation without overt evaluative pressure represents a departure from the rational assumption that force is needed to ensure program implementation.

Because the basic rational model should by now be familiar, the implementations variation will only be briefly described. In this variation of the rational model, evaluative pressure is exerted on program implementation in order to force program improvements (Leithwood and Montgomery, 1980). In turn, program improvements are assumed to lead to better educational outcomes and to the alleviation of a social problem.

But only Leithwood and Montgomery (1980) describe an evaluation of implementation in these terms. All other evaluations (or models of evaluations) of implementation either reject or ignore the use of evaluation as a means of control. For example, neither Berman and McLaughlin's (1978)

study of the implementation of innovations, nor Weiler's (1979) study of the Alum Rock voucher experiment indicated any relationship between evaluative criteria and rewards and sanctions. Similarly, evaluation models such as "contextual" (Britan, 1978a, 1978b) and "goal-free" evaluation (Scriven 1972) overtly reject pre-evaluation articulation of the evaluative criteria. Many evaluations of implementation do not attempt to exert evaluative pressure on local behavior. Thus, schools can respond to program initiatives as they wish without concern for either rewards or sanctions.

There is at least a little evidence -- though not much -- that these apparently pressureless evaluations facilitate program implementation. Alkin and Daillak (1979), for instance, report that some of the evaluators they studied developed an advisor relationship with administrators which led to program changes. Similarly, Weiler's (1979) Alum Rock study suggests that implementation may occur if evaluation does not interfere with externally stimulated change processes. These studies indicate, in effect, that evaluative pressure may interfere with implementation once schools are motivated -- usually through additional funding -- to change their programs. The absence of intense evaluative pressure may improve program implementation (Alkin and Daillak, 1979; Berman, 1978, 1980; King, Thompson and Peckman, 1981; Stern and Timar, 1981). The rational assumption that local behavior must be controlled in order to get program reforms is contradicted. Reform may occur without control.

Furthermore, Alkin and Daillak (1979) report, and Berman (1980), Cronbach and Associates (1980), and Krathwohl (1980) imply that program implementation increases when the evaluator functions more as an advisor than

as an "overseer" (Alkin and Daillak, 1979: p. 47). Trust, personal assistance, and local control of implementation, rather than the distrust, distance, and central control of implementation found in the rational model, may be fundamentally important to effective program reform (Wildavsky, 1979). Essentially, the absence of evaluative pressure may allow trusting relationships to develop that promote program implementation.

However, there are two potential problems in drawing this conclusion from these studies. First, it is possible that the evaluator is exerting evaluative pressure on implementation, even if such pressure is not overtly recognized by either the evaluator or the implementer. To illustrate this possibility, consider the somewhat farfetched example of James Popham being hired to do a goal-free evaluation. Presumably, teachers would expect to be evaluated according to a behavioral objectives framework -- even though this would not be directly communicated to them -- would want to receive good evaluations from such a well-known educator and would, as a result, attempt to teach as though they were using behavioral objectives. The presence of a well-known evaluator, especially one with clearly articulated values, is likely to exert evaluative pressure on local behaviors (Levin and Seidman, 1981). More generally, because people usually like to receive good evaluations -- even if there are no formal rewards or sanctions attached to the evaluation — they will try to conform, at least temporarily, to the evaluative criteria (Dornbusch and Scott, 1975). Thus, if a Title I evaluator visits a school, teachers will try to appear to be implementing a Title I program. Even though evaluators of implementation may appear to ignore or even overtly reject the use of evaluation as a control mechanism,

evaluative pressure focused on implementation is, nonetheless, likely to exist. This suggests that essentially low levels of evaluative pressure, rather than its absence, may actually account for the program changes.

Second, many of the studies that reported program implementation (e.g., Alkin and Daillak, 1979; Berman and McLaughlin, 1978; Weiler, 1976) were studies of projects that were perceived by local educators as both short term and peripheral to a school's instructional program. For example, the innovations examined by Berman and McLaughlin (1978) were in their last year of funding and were not widely continued. Many of the reported changes appear to be, in essence, short-term, local experimentations, that are not expected to spread throughout the system. Consequently, these types of changes are relatively non-threatening to schools and may be more easily accepted. Evaluations of implementation may only have an impact when the program to be implemented is not thought to be important.

Unfortunately, since evaluations of implementation are a relatively recent phenomena (Berman, 1978), there are as yet only a few studies of how such evaluations influence site-level behavior (King, Thompson and Peckman, 1981). Therefore, it is impossible to determine the precise impact of evaluations of implementation on local responses to reform initiatives. Nevertheless, even the limited amount of data that now exists raises questions about the necessity of using overt evaluative pressure to gain local obedience. The more subtle forms of evaluative pressure discussed above may be sufficient, in at least some circumstances, to produce program implementation, at least in peripheral programs. One of the rational model's

fundamental premises -- i.e. that local behavior must be substantially controlled in order to force local implementation -- may be incorrect.

There is, in addition, some theoretical support for this argument. Because strict evaluations are often seen as indicating distrust of local responses (Alkin and Daillak, 1979; Elmore and McLaughlin, 1981; Kagan, 1981; Shapiro, 1980), looser evaluations may indicate trust. Essentially schools may be more inclined to change if they contribute to change, if the evaluator believes they can and will change, and if they are allowed to use local discretion in determining how best to change. Also, when an evaluator functions as a negotiator (Krathwohl, 1980), educator (Cronbach and Associates, 1980) or advisor (Alkin and Daillak, 1979), the evaluator may be simultaneously exerting subtle evaluative pressure on a school and providing broader knowledge of how to implement the program. This combination may serve to both decrease the technical problems of implementation while convincing schools of the importance of the desired change. A combination of pressure, trust and assistance may be the most effective way to change schools (Berman, 1980; Elmore, 1979, 1981).

In summary, although few evaluations of implementation conform to the rational model, a limited amount of data suggests that overt control of local behavior may not be necessary to achieve program implementation. Rather, trust and advice and, perhaps, subtle evaluative pressure -- none of which are typically found in the rational model -- may be the best means of creating program reforms. The rational assumption that evaluative pressure is a prerequisite to change is problematic.

Evaluating Educational Outcomes

The assumption that evaluations of educational outcomes (Figure 2c) will cause program reform is also problematic. While focusing evaluative pressure on educational outcomes sometimes causes program improvements (Archambault and St. Pierre, 1980; Kirst, 1981; Popham, 1981), at other times evaluations of educational outcomes cause unanticipated, even counter-productive local responses (Beneviste, 1971; David, 1981; Gallagher, 1979; House, 1979; House, et al., 1978; Kagan, 1981; Levin, 1978; Madaus, 1981; McLaughlin, 1975; Nicoll, 1981; Roeber, Donovan, and Cole, 1980; Rutherford and Hoffman, 1981; Stebbins, et al., 1978; Thomas, 1981; Tyler, 1978; Weiler, 1976; Wildavsky, 1979). Again, the model will be only briefly summarized.

In this variation, educational outcomes -- typically standardized or minimum competency tests -- are evaluated as a means of forcing program change (Datta, 1981; David, 1981; Madaus, 1981; Popham, 1981). In turn, these program changes are expected to lead to improvements in the evaluated outcomes and, eventually, to alleviation of social problems such as minimal literacy among minorities. Examples of this variation of the rational model include Title I (David, 1981), the California ECE Program (Cohen and Miller, 1980), and most minimum competency tests (Popham, 1981; Tyler, 1978; Wise, 1979). In each of these programs evaluative pressure is focused on educational outcomes and is expected to lead to program improvements.

Two features of this variation require some emphasis. First, unlike the compliance variation, in which local behaviors are severely circumscribed, this variation allows at least some local autonomy. Schools are allowed to judge how best to meet the evaluative pressure. Second, although there are many different educational outcomes, e.g., self-esteem and

work habits, test scores that ostensibly measure basic skills are the most common evaluative measure (David, 1981). Therefore, this section is limited to an examination of how evaluations that use tests of basic skills affect school-level responses.

Some sources indicate that exerting evaluative pressure on test scores leads to improvements in school programs. Kirst (1981), for instance, reports that Title I evaluations have caused some improvements in schools. Similarly, Popham (1981) argues that pressure from minimum competency testing programs has forced schools to upgrade their instruction. Evaluating educational outcomes may, at times, cause desired programmatic change.

However, at other times, schools and school districts actively resist this evaluative pressure by not reporting test results (McLaughlin, 1975; Roeber, Donovan and Cole, 1980; Thomas, 1981; Weiler, 1976; Wildavsky, 1979). For example, Thomas (1981) reports that Samoan principals were ordered by their superintendent to withhold test results from federal evaluators of the Bilingual Education Act. In Alum Rock (Weiler 1976) and in Michigan (Roeber, Donovan and Cole, 1980) orders from superintendents were unnecessary since principals simply refused to supply the test data. And the resistance to reporting test data in the early years of Title I is well known (McLaughlin, 1975; Wildavsky, 1979). Some schools resist evaluation of educational outcomes by not reporting outcome data, a clear contradiction of the rational assumption that evaluation controls local behavior. Evaluation sometimes creates resistance, rather than docile behavior.

Other schools and school districts resist this evaluative pressure by sharply attacking the validity of the tests (David, 1981; Gallagher, 1979;

House, 1978; House, Glass, McLean and Walker, 1978; Levin, 1975; Madaus, 1981; Roeber, Donovan and Cole, 1980; Stebbins, et al., 1978; Tyler, 1978). Levin (1978), for instance, notes that Head Start tests were attacked as being "incomplete and biased" (page 535). Similarly, Tyler (1978) and Gallagher (1979) describe sharp attacks on the tests used in the Florida and North Carolina minimum competency testing programs. In essence, the validity of the evaluative criteria -- test scores -- is rejected. As a result, the evaluation itself is perceived as not being "soundly based" (i.e. the evaluation is considered invalid and unimportant: Dornbusch and Scott, 1975: page 95), and is therefore unlikely to cause program implementation. An invalid evaluation may not generate evaluative pressure.

In turn, because evaluations of test scores are not considered to be soundly based, schools often treat the test data as a reporting formality, decoupling these data from program decisions (David, 1981; Kagan, 1981; Madaus, 1981; McLauglin, 1975; Nicoll, 1981; Tyler, 1981). This decoupling has been reported for an elementary school in Northern California where test scores were dutifully reported then ignored (Nicoll, 1981), for many Florida schools where testing was treated as pressure for credentialling rather than as pressure for learning (Tyler, 1978) and in many Title I schools where testing was seen as another obstacle to teaching children (David, 1981; Kagan, 1981; McLaughlin, 1975). Thus, even when educational outcome data are reported, they are attacked and are decoupled from programatic decisions. The rational assumption that evaluating educational outcomes leads to program improvement is, at times, incorrect.

Furthermore, evaluating educational outcomes sometimes produces undesireable program changes. For example, several sources (Beneviste, 1981; House, 1979; Kagan, 1981; Madaus, 1981) suggest that teachers attempt to manipulate test outcomes by teaching test items rather than the skills tested. This practice improves test scores without improving students' skills, thereby creating a deceptive and potentially destructive illusion of success. In addition, Tyler (1978) reports that the threat to use test scores for evaluating teacher performance has caused teachers to switch away from difficult classes. Consequently, the classes that need good instruction the most are the least attractive classes for teachers. Evaluating educational outcomes may sometimes undermine efforts to improve educational programs. Thus, evaluating educational outcomes is, at times, counterproductive.

Several explanations for the decoupling of educational program from educational outcomes have been advanced. First, even though evaluating outcomes puts pressure on schools to change their programs, these evaluations provide no guidance about how the program should be changed (Berman and McLaughlin, 1975; David, 1981; Elmore and McLaughlin, 1981; Madaus, 1981; Sproul, forthcoming). Schools may want to improve, but may not know how to improve. Consequently, evaluating educational outcomes may have little effect on programs. Second, tests may not be valid indicators of a student's educational experience (Dickey, 1980; Elmore, 1980; Elmore and McLaughlin, 1981; House, 1979; McLaughlin, 1975). In such cases, decoupling test scores from program decisions may be a highly appropriate way to protect a program's

quality. The failures of the rational model may actually be a healthy local response.

In sum, while evaluating educational outcomes may cause educational improvements, the effects of such evaluations are, at best, inconsistent in forcing program implementation, and at worst, counterproductive to achieving a program's intent. The simplistic rational assumption that program evaluation can force desired program implementation is, at times, incorrect.

Evaluating Social Outcomes

Evaluations of social outcomes (Figure 2d) present somewhat different analytic problems than have been encountered thus far. Only two categories of programs -- vocational education and school desegregation -- consistently evaluate social outcomes, and neither of these programs conforms well to the basic rational model. However, the lack of conformity is itself an indication of the weaknesses of the rational model and is, therefore, the focus of this section.

In this variation of the rational model, it is assumed that evaluations of social outcomes such as earnings, employment or degree of school desegregation, will either directly cause program improvements or will exert pressure on educational outcomes as an indirect means of forcing program change. These program changes are expected to improve first educational outcomes -- e.g., achievement test scores -- then the evaluated social outcomes (Mertons, et al., 1980). The Vocational Educational Act of 1976 is perhaps the best example of this model since it ties a mandated evaluation of graduate employment to future funding. Evaluative pressure,

exerted on a social outcome, is expected to cause program improvement and desired social change.

Unfortunately, the literature on the implementation and evaluation of vocational education programs sheds little light on the validity of this model. This is due, at least in part, to the fact that evaluations of the social outcomes of vocational programs have only recently (in 1976) been required (Benson, 1980). Prior to that time evaluations of vocational education measured inputs. As a result, there have been relatively few evaluations of the social outcomes of vocational education, and fewer (if any) studies of the effects of such evaluations on program change. Furthermore, Mertons, et al. (1980) suggests that the short term effects of vocational education are positive enough to be above the "acceptance level" (Dornbusch and Scott, 1975: page 244) -- which is the level at which an evaluation is considered satisfactory -- of most program participants. Such evaluations are more likely to be seen as support for current program practices than as pressure f. change. Thus, with the qualification that this topic has not been well studied, the rational model appears to be irrelevant to the way vocational educational programs develop.

The literature on the implementation and evaluation of school desegregation programs, however, suggests specific inadequacies in the rational model. In particular, the role of formal evaluation as pressure for school reform -- if there is a role for formal evaluation at all -- is overshadowed both by community actions and judicial procedures (Armor, 1981; Institute for Teacher Leadership, 1979; Jensen, 1981; King, 1980; Noblit, 1979). For example, Armor's (1981) discussion of "white flight" strongly

suggests that white flight has a profound impact on schools, while Jensen's (1981) description of Judge Garrity's actions in attempting to desegregate Boston's schools indicates that the intent, bureaucratic response and evaluative pressure components of the rational model were, by virtue of judicial processes, compressed into Judge Garrity's role. In one case, community response dominates the rational model while in the other case, judicial procedures dominate the model. The rational model is, in the area of school desegregation, extremely permeable to a wide variety of external factors that decrease the importance of formal evaluative pressure.

Furthermore, there are several unexplored conceptual problems within the desegregation literature that also suggest inadequacies in the rational model. For example, there may be substantial differences in local response to pressures for integration as opposed to desegregation (King 1980) and to required busing versus required magnet schools (Levine and Havighurst, 1977). Unfortunately, while there are a huge number of studies of desegregation, few have studied the effects of different desegregation schemes on school programs (Jensen, 1981; King, 1980; Levine and Havighurst, 1977). But the very existence of these distinctions suggests that the simplicity of the rational model significantly underestimates the complexity of school desegregation programs. The validity of the rational model is suspect.

Several arguments have been advanced to explain both the paucity of programs that evaluate social outcomes and the inadequacies of this variation of the rational model. First, as noted earlier, it is often quite difficult to translate vague social goals into valid, operational measures of social outcomes (Ashenfelter, 1980; David, 1981; Elmore and McLaughlin, 1981; House,

1979; Levin, 1978; Martini and Associates, 1972; Wildavsky, 1979). House (1979), for instance, reports that SRI made substantial efforts to operationalize the social outcomes of Follow Through, but ultimately abandoned the effort as impossible. Because valid measures of social outcomes are difficult to obtain, evaluators may shy away from measuring such outcomes and educators may reject such measures as unsoundly based. Second, even if a measure of a social outcome is valid, it may be affected by so many other factors (e.g., regional unemployment or latent racism) that the effects of a particular program on the outcome can not be specifically determined (Armor, 1981; Ashenfelter, 1980; Benson, 1980; Levin, 1978; Martini and Associates, 1972; Noblit, 1979). This is clearly the case for desegregation since community actions, judicial processes or many other factors may, at times, interact with and overwhelm program effects. Such circumstances are likely to lead to a belief that an evaluation is not soundly based and can, therefore, be ignored. Third, the social outcomes of a program may not appear until many years after a student has completed the program (Levin, 1978). The effects of Head Start on minority unemployment and poverty might not be apparent for more than a decade after children have completed the program. Because a substantial amount of time may elapse before a program's social outcomes can be measured, short-term measures of these outcomes are likely to be seen as inappropriate while long term measures are unlikely to affect a program. Evaluations of a program's social outcomes may be, in effect, temporally decoupled from the program. Given such circumstances, evaluating social outcomes is unlikely to lead to program reform.

In summary, the rational assumption that evaluations of social outcomes produce program implementations appears questionable. Evaluations of social outcomes are irrelevant to change in vocational education programs while the various external factors that affect school desegregation suggest that the presumed efficiency of the rational model is too simplistic. Essentially, simplistic rational assumptions are overwhelmed by a complex reality.

Mixed Models

Although the rational assumption that evaluative pressure can change behavior has been decomposed into four variations, each of which was examined separately, this is not a completely accurate representation of the rational model. Actually, many programs have evaluations that focus on several different components simultaneously. Title I, for example, mandates both a compliance and educational outcomes evaluation, while PL94-142 requires both a compliance and an implementation evaluation. Many programs are combinations of the variations examined above. However, there has been little or no examination of either the combined or interactive effects of exerting evaluative pressure on several different components, so this topic will not be examined further.

Evaluation and Control

This analysis strongly suggests that the rational assumption that evaluative pressure can control local behaviors and is necessary to force program implementation is, at times, incorrect. Evaluative pressure exerted on compliance produces compliance, but compliance does not always produce implementation. Evaluative pressure is not usually exerted on implementation, but program reforms sometimes occur anyway. Evaluative

pressure exerted on educational outcomes produces defensive reactions but does not necessarily produce program reform. Evaluative pressure is rarely exerted on social outcomes, but, when such pressure is exerted, its effects are overwhelmed by larger social factors. In sum, evaluative pressure does not produce the anticipated local responses.

Consequently, decision makers and evaluators, operating on the assumption that evaluation controls behavior, are confused and frustrated by their inability to impose change on schools. In their frame of reference, evaluation that does not produce control is impossible. Thus, the public blames ineffective programs, decision makers accuse evaluators of examining the wrong variables, and evaluators accuse decision makers of ignoring the evaluation results. The unanticipated wealmesses of the rational model leads to growing dissatisfaction with program implementation and evaluation.

It is important to emphasize, however, that resistance to rational program implementation and evaluation may be an appropriate response to unreasonable pressures. Mass programs may not be suitable for any particular student or school. Furthermore, government intervention into schools may be a threat to democracy since it increases centralized control of an important socializing institution. While there are some areas where government intervention into the educational system seems legitimate -- e.g., school desegregation -- local rejection of other government interventions may be a healthy response to inappropriate reform initiatives. Recognition of appropriate versus inappropriate resistance to the rational model may

decrease dissatisfaction with the presumed inadequacies of educational reforms.

Summary and Implications

As may be apparent, the argument advanced thus far is an extensive attack on the rational model of program implementation and evaluation. The assumed linkage between program intent and evaluative criteria is problematic, as is the assumption that program evaluations can force or are necessary for program implementation. In a sense, the two most crucial aspects of a social program, intent and implementation, may not be a part of most programs and program evaluations. This possibility has many research and practical implications.

Research

At least three research implications can be identified. First, further research is needed on the linkage between program intent and evaluative criteria. Although it is probable that this linkage is weak, the exact nature of this linkage has not been empirically determined. Second, while the effects of evaluating compliance on school behaviors have been relatively well explored, the effects of evaluating implementation, education and social outcomes or combined variations, have received little attention. Comparisons of the different variations may be a fruitful avenue of research since these might indicate which of the evaluative models produced the most appropriate change in schools. Third, because evaluations exert pressure on schools, schools that are evaluated are likely to respond to program initiatives differently than schools that are not evaluated. As a result, data taken from a sample of schools is generalizeable only to other schools

in the sample, and not to schools that have not been evaluated. Evaluations, in effect, contaminate their own data. Verifying or refuting this last implication is a substantial and important research project in its own right. Thus, this analysis of the rational model suggests a large and potentially valuable research agenda.

Practical

This paper highlights one of the central dilemmas in education today. While the history of school desegregation and accountability suggests that schools should not be entirely trusted to implement needed reforms, government programs that rely on evaluative pressure have not always produced the desired results. How, then, can reluctant schools be made to respond to legitimate social and educational needs, without having government impose unrealistic, inappropriate, and, ultimately, unsuccessful demands on local educators?

Unfortunately, there is probably no single "good" answer to this question. Rather, the best response seems to be a sort of facilitator role, in which a government agent -- the evaluator -- both exerts evaluative pressure on a school and acts as an advisor on how to integrate the new program into on-going activities. Presumably, at times, the facilitator will decide that a program is appropriate for a district and support local initiatives while, at other times, the facilitator will decide that a program is inappropriate for a district -- either because the district has other legitimate priorities, or because the district fails to implement the program -- and will terminate the program. Such a role combines elements of Alkin and Daillak's (1979) "adviser," Berman's (1980) "adaptive implementation,"

Cronbach and Associates (1980) "educator" and Krathwohl's (1980) "negotiator-fact finder." The evaluator both pushes for and assists program reform.

However, there are some problems with this role conception. Implementation may rely very heavily on the evaluator's judgement and social skills, neither of which will necessarily be of high quality. Also, since a personal relationship must be established at the school level, this model is quite labor intensive, and therefore, quite expensive. Nevertheless, the evaluator as facilitator model appears to be the most promising approach to producing an increase in program implementation, an increase in the usefulness of program evaluations, and a decrease in the current dissatisfaction with large-scale educational programs and program evaluation.

REFERENCES

Alkin, Marvin C. and Richard H. Daillak. "A Study of Evaluation Utilization." Educational Evaluation and Policy Analysis Vol. 1, No. 4 (July-August, 1979): p. 41-49.

Alkin, Marvin C., Richard Daillak, and Peter White. Using Evaluations. Beverly Hills: Sage, 1979.

Archambault, Francis X., and Robert G. St. Pierre. "The Effect of Federal Policy on Services Delivered Through ESEA Title I." Educational Evaluation and Policy Analysis Vol. 2, No. 3 (May-June, 1980): p. 33-46.

Armor, David J. "White Flight and the Future of School Desegregation." in Evaluation Studies Review Annual Vol. 5, 1981, Howard E. Freeman and Marian A. Soloman (eds). Beverly Hills: Sage, 1981.

Ashenfelter, Orley. "Estimating the Effect of Training Programs on Earnings." in Evaluation Studies Review Annual Vol. 5, 1980, Ernst W. Stromsdorfer and George Farkus (eds). Beverly Hills: Sage, 1980.

Benson, Charles S. Centralization and Legalization in Vocational Education: Limits and Possibilities, Program Report No. 80-B20, Institute for Research on Educational Finance and Governance, Stanford University, Stanford, CA, December, 1980.

Benveniste, Guy. Implementation and Intervention Strategies: The Case of PL94-142. Project Report No. 81-A18, Institute for Research on Educational Finance and Governance, Stanford University, Stanford, CA, September, 1981.

Berman, Paul. "The Study of Macro-and Micro-Implementation." Public Policy Vol. XXVI, No. 2 (Spring 1978): p. 157-84.

Berman, Paul. "Thinking About Programmed and Adaptive Implementation: Matching Strategies to Situations." in Why Policies Succeed or Fail, Helen M. Ingram and Dean E. Mann (eds). Beverly Hills: Sage, 1980: pp. 205-227.

Berman, Paul. "From Compliance to Learning: Implementing Legally-Induced Reform." Project Report No. 81-A20, Institute for Research on Educational Finance and Governance, Stanford University, Stanford, CA, December, 1981.

Berman, Paul, and Milbrey W. McLaughlin. Federal Programs Supporting Educational Change, Vol VIII: Implementing and Sustaining Innovations. R-1589/8-HEW. Santa Monica: Rand, May, 1978.

Britain, Gerald M. "The Place of Anthropology in Program Evaluation." Anthropology Quarterly Vol. 51, No. 2 (April, 1978a) pp. 119-128.

Britain, Gerald M. "Experimental and Contextual Models of Program Evaluation." in Education and Program Planning, Vol. 1. Pergamon Press, 1978b, p. 229-234.

Cohen, Elizabeth G., and Russell H. Miller. "Coordination and Control of Instruction in Schools." Pacific Sociological Review Vol. 23, No. 4. (October, 1980): pp. 446-473.

Cronbach, Lee J., and Associates. Toward Reform of Program Evaluation. San Francisco: Jossey-Bass, 1981.

Cyert, Richard M., and James G. March. A Behavioral Theory of the Firm. Englewood Cliffs, NJ: Prentice-Hall, 1963.

Datta, Lois-ellin. "Damn the Experts and Full Speed Ahead." Evaluation Review Vol. 5, No. 1 (February, 1981): pp. 5-32.

David, Jane L. "Local Uses of Title I Evaluations." Educational Evaluation and Policy Analysis Vol. 3, No. 1 (January-February, 1981): pp. 27-39.

Dickey, Barbara. "Utilization of Evaluations of Small-Scale Innovative Educational Projects." Educational Evaluations and Policy Analysis Vol. 2, No. 6 (November-December, 1980): pp. 65-77.

Dornbusch, Sanford M., and Richard W. Scott. Evaluation and the Exercise of Authority. San Francisco: Jossey-Bass, 1975.

Elmore, Richard F. "Complexity and Control: What Legislators Can Do About Policy Implementation." Seattle: Institute for Public Policy Research (Paper No. 11), University of Washington, 1979.

Elmore, Richard F. "Evaluation, Control, and Learning in Organizations." A paper presented at the Annual Meeting of the Western Political Science Association. San Francisco: 1980.

Elmore, Richard F. "Implementation and the Law: Authority, Responsibility, and Competence." A paper presented at the Annual Meeting of the Association for Public Policy Analysis and Management. Washington: 1981.

Elmore, Richard F., and Milbrey W. McLaughlin. "Strategic Choice in Federal Education Policy: The Compliance-Assistance Trade-Off." Draft, June, 1981.

Felix, Joseph L. "Research and Evaluation to Improve Instruction: The Cincinatti Strategy." Educational Evaluation and Policy Analysis Vol. 1, No. 2 (March-April, 1975): pp. 57-62.

Gallagher, James J. "Minimum Competency: The Setting of Educational Standards." Educational Evaluation and Policy Analysis Vol. 1, No. 1 (January-February, 1979): pp. 62-67.

Gideonse, Hendrik D. "Improving the Federal Administration of Educational Programs." Educational Evaluation and Policy Analysis Vol. 2, No. 1 (January-February, 1980): pp. 61-70.

Hoachlander, E. Gareth. Distributing Federal Categorical Aid by Formula: The Case of Vocational Education. Project Report No. 81-A22. Institute for Research on Educational Finance and Governance, Stanford University, Stanford, CA: October, 1981.

House, Ernest R. "The Objectivity, Fairness, and Justice of Federal Evaluation Policy As Reflected in the Follow Through Evaluation." Educational Evaluation and Policy Analysis Vol. 1, No. 1 (January-February, 1979): pp. 28-42.

House, Ernest R., Gene V. Glass, Leslie D. Mclean, and Decker F. Walker. "No Simple Answer: Critique of the Follow Through Evaluation." in Evaluation Studies Review Annual Vol. 3, 1978, Richard D. Look, et al. (eds). Beverly Hills: Sage, 1978.

Institute for Teacher Leadership. The Prevention of Resegregation: Strategies for Teachers. Washington: DHEW, July, 1979.

Jensen, Donald N. Desegregation in Boston: The Federal Courts as School Administrator. Program Report No. 81-B12, Institute for Research on Educational Finance and Governance, Stanford University, Stanford, CA: November, 1981.

Kagan, Robert A. "Regulating Business Regulating Schools: The Problem of Regulatory Unreasonableness." Project Report No. 81-A14. A paper presented at the IFG Seminar on Law and Governance in Education. Stanford, CA: July, 1981.

King, A. L. Executive Summary of a Final Report on Analysis of Staff Development Programs and Other Strategies to Improve Education in Desegregated Schools. Austin, TX: Southwest Educational Development Laboratory, 1980.

King, Jean A., Bruce Thompson, and Ellen M. Pechman. Improving Evaluation Use in Local School Settings. New Orleans: New Orleans Public Schools, June, 1981.

Kirst, Michael. "Teaching Policy and Federal Categorical Programs." A paper prepared for the National Institute of Education Program on Teaching Policy, Spring, 1981.

Kirst, Michael, and Richard Jung. "The Utility of a Longitudinal Approach in Assessing Implementation: A Thirteen Year View of Title I, ESEA." Program Report No. 80-B18, Institute for Research on Educational Finance and Governance, Stanford University, Stanford, CA: June, 1980.

Krathwohl, David R. "The Evaluator as Negotiations Facilitator-Fact Finder." Educational Evaluation and Policy Analysis Vol. 2, No. 2 (March-April, 1980): pp. 25-34.

Leithwood, Kenneth A., and Deborah J. Montgomery. "Evaluating Program Implementation." Evaluation Review Vol 4, No. 2 (April, 1980): pp. 193-214.

Levin, Henry M. "Economic Evaluation of Education." in Speaking Out: Law, Education, and Politics. Princeton, NJ: Evaluation Testing Service, 1974, pp. 84-96.

Levin, Henry M. "A Decade of Policy Developments in Improving Education and Training for Low Income Populations." in Evaluation Studies Review Annual, Vol. 3, 1978. Thomas D. Cook et al. (eds). Beverly Hills: Sage, 1978.

Levin, Henry M., and William Seidman. "Towards a Meta Cost-Effectiveness Analysis of Educational Interventions." Program Report No. 81-B13. Institute for Research on Educational Finance and Governance, Stanford University, Stanford, CA: December, 1981.

Levine, Daniel U., and Robert J. Havighurst. "Introduction, the Desegregation Controversy in Context." in The Future of Big-City Schools Daniel U. Levine and Robert J. Havighurst (eds). Berkeley: McCutchan, 1977.

Madaus, George F. "NIE Clarification Hearing: The Negative Team's Case." Phi Delta Kappan Vol. 63, No. 2 (October, 1981): pp. 92-94.

Majone, Giandomenico. "Process and Outcome in Regulatory Decision-Making." in Making Bureaucracies Work, Carol H. Weiss and Allen H. Barton (eds). Beverly Hills: Sage, 1980.

March, James G., and Johan P. Olsen. Ambiguity and Choice in Organizations Bergen, Norway: Universitetsforlajet, 1976.

Martini and Associates. Vocational and Special Education. Anchorage, Alaska: Martini and Associates, March, 1972.

McLaughlin, Milbrey W. Evaluation and Reform. Cambridge, MA: Ballinger, 1975.

Mertens, et al. The Effects of Participating in Vocational Education: Summary of Studies Reported Since 1968. Columbus, OH: The National Center for Research in Vocational Education, May, 1980.

Meyer, John W., and Brian Rowan. "Institutionalized Organizations: Formal Structure Versus Myth and Ceremony." American Journal of Sociology Vol. 83, No. 2 (1977): pp. 340-363.

Meyer, John W., and Brian Rowan. "The Structure of Educational Organizations." in Environments and Organizations, Marshall W. Meyer and Associates (eds). San Francisco: Jossey-Bass, 1978.

Monti, Daniel J. "Administrative Discrimination in the Implementation of Desegregation Policies." Educational Evaluation and Policy Analysis Vol. 1, No. 4 (July-August, 1979): pp. 17-25.

National Institute of Education. Administration of Compensatory Education. Washington, DC: 1977.

Nicoll, Paul. "School District Evaluation of a Northern California Elementary School." Evaluation Review Vol. 5, No. 2 (June, 1981): pp. 409-424.

Noblit, George W. Managing Student Desegregation: A Comparative Study of the Functioning of Principals and Implications for Policy Research. Washington, DC: DHEW, September, 1979.

Pfeffer, Jeffrey, and Gerald R. Salancik. The External Control of Organizations. New York: Harper and Row, 1978.

Popham, W. James. "The Case for Minimum Competency Testing." Phi Delta Kappan, Vol. 63, No. 2 (October, 1981): pp. 89-91.

Roeber, Edward D., David L. Donovan, and Richard T. Cole. "Telling the Statewide Testing Story . . . and Living to Tell It Again." Phi Delta Kappan, Vol. 62, No. 4 (December, 1980): p. 273.

Rutherford, William L., and James V. Hoffman. "Toward Implementation of the ESEA Title I Evaluation and Reporting System: A Concerns Analysis." Educational Evaluation and Policy Analysis, Vol. 3, No. 4 (July-August, 1981): pp. 17-23.

Sabatier, Paul, and Daniel Mazmanian. "The Implementation of Regulatory Policy: A Framework of Analysis." Paper No. 39, Institute for Government Affairs, Davis, CA: n.d.

Scott, Richard W. Organizations. Englewood Cliffs, NJ: Prentice-Hall, 1981.

Scriven, Michael. "Prose and Cons About Goal-Free Evaluation." Evaluation Comment Vol. 3, No. 4 (December, 1972): pp. 1-4.

Shapiro, Martin. Federal Regulation of Education. Program Report No. 80-B17. Institute for Research on Educational Finance and Governance, Stanford University, Stanford, CA: July, 1980.

Sproul, Lee S. "Responding to Regulations: School Superintendents and the Federal Government." in Organizational Analysis of Schools and School Districts, Samuel Bachrack (ed). New York: Praeger, 1981a.

Sproul, Lee S. "Response to Regulation: An Organizational Process Framework." Administration and Society, Vol. 12, No. 4 (February, 1981b): pp. 447-470.

Sproul, Lee S. "Managing Education Programs: A Micro-Behavioral Analysis." in Human Organization, forthcoming.

Stearns, Marian S., David Greene, and Jane L. David. Local Implementation of PL94-142: First Year Report of a Longitudinal Study. Stanford Research Institute, Menlo Park, CA: April, 1980.

Stebbins, Lind B., et al. "An Evaluation of Follow Through." in Evaluation Studies Review Annual, Vol. 3, 1978, Thomas D. Look, et al. (eds). Beverly Hills: Sage, 1978.

Stern, David, and Thomas Timar. Conflict and Choice in Public Education. Program Report No. 81-A21, Institute for Research on Educational Finance and Governance, Stanford University, Stanford, CA: November, 1981.

Talbert, Joan E. "School Organization and Institutional Change: Exchange and Power in Loosely Coupled Systems." Project Report No. 80-A9, Institute for Research on Educational Finance and Governance, Stanford University, Stanford, CA: August, 1980.

Thomas, R. Murray. "Evaluation Consequences of Unreasonable Goals - The Plight of Education in American Samoa." Educational Evaluation and Policy Analysis, Vol. 3, No. 2 (March-April, 1981): pp. 41-49.

Tyler, Ralph W., et al. The Florida Accountability Program. Washington, DC: National Education Association, July, 1978.

Weatherley, Richard, and Michael Lipsky. "Street-Level Bureaucrats and Institutional Innovation: Implementing Special Education Reform." Harvard Educational Review, Vol. 47, No. 2 (May, 1977): pp. 171-197.

Weick, Karl E. "Educational Organizations as Loosely Coupled Systems." Administrative Science Quarterly, Vol. 21 (March, 1976): pp. 1-19.

Weick, Karl E. The Social Psychology of Organizing. Reading, MA: Addison-Wesley, 1979.

Weiler, Daniel. "A Public School Voucher Demonstration: The First Year of Alum Rock Summary and Conclusions." in Evaluation Studies Review Annual, Vol. 1, 1976, Gene V. Glass (ed). Beverly Hills: Sage, 1976.

Wholey, Joseph S. "Using Evaluation to Improve Program Performance." in Evaluation Studies Annual Review, Vol. 6, 1981. Howard E. Freeman and Marian A. Solomon (eds). Beverly Hills: Sage, 1981.

Wildavsky, Aaron. Speaking Truth to Power. Boston: Little, Brown, and Co., 1979.

Wilkes, Christopher D., Vicki Lynn Eaton, Kenneth D. Wood, and Sanford M. Dornbusch. "Centralization and Tighter Coupling in Education Through Program Evaluation." Project Report No. 79-A6, Institute for Research on Educational Finance and Governance, Stanford University, Stanford, CA: September, 1979.

Wise, Arthur E. Legislated Learning. Berkeley: University of California Press, 1979.

Figure 1: The Rational Model of Program Implementation and Evaluation

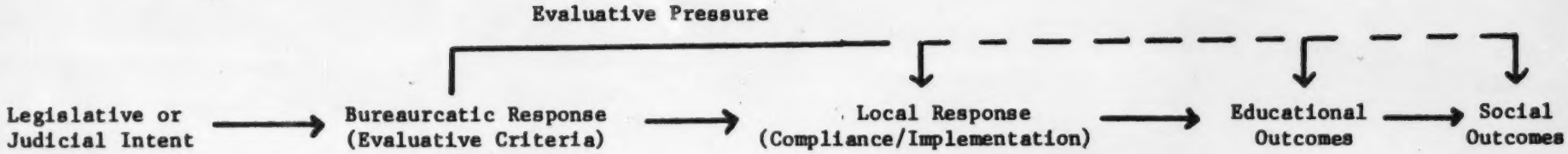

Figure 2: Variations on the Rational Assumption That Program Evaluation Controls Site-Level Behaviors